European Journal of Protistology 91 (2023) 126013

Contents lists available at ScienceDirect

European Journal of Protistology

journal homepage: www.elsevier.com/locate/ejop

# Expansion of the cytochrome C oxidase subunit I database and description of four new lobose testate amoebae species (Amoebozoa; Arcellinida)

Giulia M. Ribeiro [a], Fernando Useros [b], Kenneth Dumack [c], Rubén González-Miguéns [b], Ferry Siemensma [d], Alfredo L. Porfírio-Sousa [a], Carmen Soler-Zamora [b], João Pedro Barbosa Alcino [a], Daniel J.G. Lahr [a], Enrique Lara [b,*]

[a] Department of Zoology, Biosciences Institute, University of São Paulo, São Paulo, Brazil
[b] Real Jardín Botánico (RJB-CSIC), Madrid, Spain
[c] Department of Terrestrial Ecology, Institute of Zoology, University of Cologne, Germany
[d] Julianaweg 10, 1241VW Kortenhoef, Netherlands

ARTICLE INFO

*Keywords:*
Arcellinida
Taxonomy
Next generation sequencing
Bioinformatics
Data mining

ABSTRACT

Arcellinida is ascending in importance in protistology, but description of their diversity still presents multiple challenges. Furthermore, applicable tools for surveillance of these organisms are still in developing stages. Importantly, a good database that sets a correspondence between molecular barcodes and species morphology is lacking. Cytochrome oxidase (COI) has been suggested as the most relevant marker for species discrimination in Arcellinida. However, some major groups of Arcellinida are still lacking a COI sequence. Here we expand the database of COI marker sequences for Arcellinids, using single-cell PCR, transcriptomics, and database scavenging. In the present work, we added 24 new Arcellinida COI sequences to the database, covering all unsampled infra- and suborders. Additionally, we added six new SSUrRNA sequences and described four new species using morphological, morphometrical, and molecular evidence: *Heleopera steppica, Centropyxis blatta, Arcella uspiensis,* and *Cylindrifflugia periurbana*. This new database will provide a new starting point to address new research questions from shell evolution, biogeography, and systematics of arcellinids.

## 1. Introduction

Arcellinida or lobose testate amoebae are increasingly studied in a wide array of applied and fundamental disciplines, including ecology, evolution, bioindication, biogeography and paleoecology (Arrieira et al., 2015, Atasiei et al., 2022, González-Miguéns et al., 2022b, Heger et al., 2013, Marcisz et al., 2020, Nasser et al., 2020, Soler-Zamora et al., 2021). Estimations based on bibliometrics show that the importance of testate amoebae in all research areas increased significantly during the past decade (Kosakyan et al., 2016). All these findings are facilitated by the development of an updated and better-resolved taxonomic framework, which is currently being refined by the use of phylogenetics and transcriptomics (Lahr et al., 2019) and by an increased taxon sampling (González-Miguéns et al., 2022a). Currently, the order Arcellinida is divided into three suborders, Phryganellina, Organoconcha, and Glutinoconcha (Lahr et al., 2019), the latter comprising at least six infraorders whose number may increase in the future with the addition of some orphan taxa ("*incertae sedis*", González-Miguéns et al. (2022b)). But while it is unlikely that the backbone of the Arcellinida tree may change dramatically in the next few years, the evaluation of species-level biodiversity is still challenging.

There are only a few hundreds of Arcellinida species described nominally (Meisterfeld, 2001; Kosakyan et al., 2016). The traits that are used to distinguish these species are traditionally to be found on the organisms shells,including shape/size, composition and ornamentations (Kosakyan et al., 2016). However, the correspondence between these traits and species diversity is not always obvious. Indeed, shell morphology varies to a certain extent with environmental conditions (Bobrov and Mazei, 2004, Medioli et al., 1987, Porfírio-Sousa and Ribeiro et al., 2017, Schulz et al., 2018, Mulot et al., 2017). Conversely, there is a large amount of cryptic diversity, meaning that different species may have identical morphology (Kosakyan et al., 2012). There is also evidence of evolutionary convergences, where genetically far-related species that live in comparable environments can exhibit

* Corresponding author at: Real Jardín Botánico (RJB-CSIC), Plaza Murillo 2, 28014 Madrid, Spain.
*E-mail address:* enrique.lara@rjb.csic.es (E. Lara).

https://doi.org/10.1016/j.ejop.2023.126013

Available online 13 August 2023

similar morphologies (Lahr et al 2014, González Miguens 2022a, 2022b). For all these reasons, shell morphology alone is not a reliable proxy for species identification. Species are the basic unit of biodiversity (Leakey and Lewin, 1996; Wiley, 1978), on which ecological and/or evolutionary hypotheses are to be tested. Therefore, if species characterization remains unreliable, experimental results are biased and conclusions may be misleading.

Molecular taxonomic identifications have been increasingly proposed as a way to overcome morphological identification issues (Kosakyan et al., 2012, Macumber et al., 2020, Pawlowski et al., 2013). For that purpose, a comprehensive database that establishes congruence between sequences and morphology is needed; in other terms, the barcoding approach (Hebert et al., 2003; Pawlowski et al. 2012). Nowadays, there are mainly three molecular markers that are used to build phylogenies and evaluate Arcellinida diversity; the nuclear SSU rRNA (coding for the small subunit of the ribosomal RNA), the mitochondrial Cytochrome Oxidase Subunit I (COI) and the Nicotinamide Adenine Dinucleotide Dehydrogenase (NAD) (Blandenier et al., 2017; González-Miguéns et al., 2022b). While mitochondrial markers have a taxonomic resolution to the species level, SSU rRNA often fails in delimiting species (Lara et al., 2008) but is suited for deep node reconstruction in Arcellinids (González-Miguéns et al., 2022b). For practical reasons such as easier amplification, COI has become the most commonly applied marker in Arcellinida; it is also the one for which the largest database is available (i.e. 687 entries in National Center for Biotechnology Information, NCBI database by September 2022). However, in order to make inferences about Arcellinida environmental diversity, it remains necessary to build a database which establishes the correspondence between COI sequences and shell morphologies (Jones, 2008).

Presently, the representation of Arcellinida COI sequences in public databases is taxonomically biased. Indeed, most of the studies based on COI barcodes addressed Hyalospheniformes, which represent around 90% (610 entries by September 2022) of COI data in National Center for Biotechnology Information (NCBI). On the other hand, the infraorder Volnustoma and the still unplaced genus *Argynnia* (Glutinoconcha) have still not been sampled for COI; the entire suborders Organoconcha and Phryganellina are not represented. In this work, our aim is to increase the Arcellinida molecular database of the barcoding genes. In that purpose, we (1) isolated and barcoded single cells, using COI and SSU genes, from underrepresented Arcellinida taxa (2) explored for COI and SSU genes in different transcriptomic surveys, making them available in public repositories such as GenBank and (3) sequenced transcriptomes from *Phryganella acropodia*. Furthermore, we add the first COI data from the infraorders Volnustoma, Organoconcha and Phryganellina. Finally, we describe four new species, *Heleopera steppica* sp. nov., *Centropyxis blatta* sp. nov., *Arcella uspiensis* sp. nov. (previously *Arcella intermedia*), and *Cylindrifflugia periurbana* sp. nov. (previously *Cylindrifflugia acuminata*), on the basis of ecological, molecular and morphological data.

## 2. Material and methods

- Sampling and microscopical observation of Arcellinids specimens

We aimed at isolating specimens from undersampled groups (i.e. genera *Arcella, Argynnia, Centropyxis, Cylindrifflugia, Difflugia, Heleopera,* and *Lagenodifflugia*), which were obtained from freshwater sediment or terrestrial mosses. For each freshwater sample, we took only the uppermost few millimetres on top of the sediment with a plastic Pasteur pipette and deposited it in a falcon tube with water from the sampled water body. Moss samples were humidified and squeezed in a petri dish to extract active arcellinids. Individual cells were collected and photodocumented under an inverted microscope (Leica DMI8, 400x magnification) using a stirred Pasteur pipette. Some organisms could be maintained either as monoprotistan or as enrichment cultures. For the cultures, we used a growth medium that consists of autoclaved water taken from the site where they were sampled and cereal grass media (Cerophyll - Fisher Scientific S25242) as described previously (Porfírio-Sousa and Ribeiro et al., 2017). For those maintained in enrichment cultures, several (6 to 10) cells were isolated and placed in a petri dish containing water from the collection site plus an autoclaved rice grain. We collected a different number of individuals for each lineage, depending on the availability in the environment or the cultures (Table 1).

The species that would be subsequently described (*Heleopera steppica, Arcella uspiensis, Centropyxis blatta* and *Cylindrifflugia periurbana*), were also documented based on SEM microscopy and morphometrical analysis. For that purpose, we placed individuals on stubs for electron microscopy; these stubs will constitute the type material. The SEM observations took place at the facility of the Real Jardín Botánico de Madrid as described previously (González-Miguens et al., 2022). For morphometrical analysis, we measured shell length, width, and aperture size (i.e. " traditional measurements") to enable comparisons with the original literature (Fig. 1). We used ImageJ (v.1.52) (Schneider et al., 2012) to capture the measurements directly from the pictures. All specimens were at first identified morphologically using original descriptions. Later, the identifications were corrected if necessary using the result of the phylogeny.

- Molecular work

We followed a standard protocol based on guanidine buffer extraction (Duckert et al., 2018), where we isolated single cells and cleaned them in a series of Milli-Q filtered water. We then put the clean cell into 100 µl of guanidine and performed the extraction according to the protocol. Alternatively, we placed the cell directly in a PCR tube containing the PCR mix. We amplified a fragment of the mitochondrial cytochrome oxidase (COI) from DNA extractions and also SSU, whenever possible. For the COI amplification, we used a two-step PCR approach as previously described (González-Miguéns et al., 2022a; 2023). The first round of PCR relied on an amplification using the universal primers LCO 1490 (5′GGTCAACAAATCATAAAGATATTGG3′) and HCO 2198 (5′TAAACTTCAGGGTGACCAAAAAATCA3′) (Folmer et al., 1994). In the second round of PCR, we applied a semi-nested approach (González-Miguéns et al., 2022a; 2023) which combined the Arcellinida-specific primer ArCOIR (5′CCACYNGAATGWGCTARAAT ACC3′) and the broad spectrum LCO. For SSU, we used the universal eukaryotic primers EK555F (5′AGTCTGGTGCCAGCAGCCGC3′) and EK1498R (5′ CACCTACGGAAACCTTGTTA3′) as described previously (González-Miguéns et al., 2022b). Bands of the right size were excised from an agarose gel and sequenced at Macrogen Inc. (Macrogen Europe, Madrid, Spain).

The first isolation of *Arcella* sp. then named *Arcella intermedia,* took place in 2013. This organism has been maintained in monoclonal culture since then and was used in several transcriptomic, morphometric, and competition studies (Lahr et al., 2019, Porfírio-Sousa and Ribeiro et al., 2017, Ribeiro et al., 2020, 2019). However, culture vitality declined strongly since then, and the organism had to be isolated again from the same locality, a small artificial pond within the University of São Paulo (Coordinates –23.565720, −46.730512). The new isolate has the same morphology and COI barcode as the previous one and thus both can be considered as conspecific.

*Phryganella paradoxa* (strain TP) and *Phryganella acropodia* (strain 3112, 31111, and 3115) derive from monoclonal cultures grown with diatoms as food (*Synedra* sp. strain CCAC 1762B and *Nitzschia communis* strain CCAC 1762) as described previously (Dumack et al., 2019). We extracted the RNA from single cells using smart-seq2 protocol (Picelli et al., 2014; Lahr et al., 2019). The RNA was then reverse-transcripted, and we prepared a library for subsequent sequencing using the kit Nextera-XT. We sequenced the cDNA using an Illumina NextSeq platform (outsourced at Centro de Facilidades de Apoio a Pesquisa, CEFAP-USP). We first filtered and removed adaptors from the raw reads file using Bbtools (http://sourceforge.net/projects/bbmap) with a

**Table 1**
New sequences from phylogenetic markers COI and SSUrDNA added to the Arcellinids database. In this work, we added 46 new sequences to the Arcellinids database. 23 from COI marker an 23 from SSUrDNA marker. All the information about the taxon are included here: classification, data source, locality of sampling an type of information available. Column names code: OM – optical microscopy, SEM – electron microscopy, Ma- morphometric analysis.

| Taxon | Group | Source | Country | Locality | Molecular Data | OM | SEM | Ma | N |
|---|---|---|---|---|---|---|---|---|---|
| *Arcella uspiensis* Sp. Nov. | Sphaerothecina | Monoclonal Culture/ Database Search | Brazil | –23.565720, −46.730512 | Transcriptome SSU (PRJNA513164, PRJNA515423)/ Single-Cell Barcode COI (OQ631082) | Yes | Yes | Yes | 52 |
| *Arcella* Sp. LKH685 | Sphaerothecina | Database Search | USA | 42.315406, −72.642303 | Transcriptome (PRJNA846601) | No | No | No | – |
| *Netzelia Tuberspinifera* C15_1 | Sphaerothecina | Database Search | China | 24.800000, 118.133333 | Transcriptome (PRJNA698229) | No | No | No | – |
| *Netzelia Tuberspinifera* C15_2 | Sphaerothecina | Database Search | China | 24.800000, 118.133333 | Transcriptome (PRJNA698229) | No | No | No | – |
| *Netzelia* Sp. | Sphaerothecina | Database Search | Brazil | –22.677778, −42.974444 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Cyclopyxis Lobostoma* | Sphaerothecina | Database Search | Brazil | –23.565708, −46.728689 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Centropyxis blatta* Sp. Nov. | Excentrostoma | Cultive | Spain | 40.229,-3.758 | Transcriptome SSU (PRJNA945922)/ Single-Cell Barcode COI (OQ631081) | Yes | Yes | Yes | 25 |
| *Centropyxis* sp. | Excentrostoma | Database Search | Brazil | –23.566339, −46.730632 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Centropyxis aculeata* | Excentrostoma | Database Search | Brazil | –22.677778, −42.974444 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Cylindrifflugia periurbana* Sp. Nov. | Cylindrothecina | Environmental | Spain | 40.229,-3.758 | Single-Cell Barcode (OL549144, OL549145, OL549146) | Yes | Yes | Yes | 4 |
| *Heleopera steppica* Sp. Nov. | Volnustoma | Environmental | Spain | 39.991667, −3.615278 | Single-Cell Barcode COI (OQ631086), SSUrDNA (OQ646056) | Yes | Yes | Yes | 25 |
| *Heleopera sylvatica* | Volnustoma | Database Search | Brazil | –23.566339, −46.730632 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Heleopera sphagni* | Volnustoma | Database Search | France | 46.831247, 6.160270 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Phryganella acropodia* Pacr31111 | Phryganellina | Monoclonal Culture | Germany | – | Transcriptome (PRJNA945922) | No | No | No | – |
| *Phryganella acropodia* Pacr3115 | Phryganellina | Monoclonal Culture | Germany | – | Transcriptome (PRJNA945922) | No | No | No | – |
| *Phryganella acropodia* Pacr3112 | Phryganellina | Monoclonal Culture | Germany | – | Transcriptome (PRJNA945922) | No | No | No | – |
| *Phryganella paradoxa* TP | Phryganellina | Monoclonal Culture | Germany | – | Transcriptome (PRJNA945922) | No | No | No | – |
| *Cryptodifflugia operculata* | Phryganellina | Database Search | USA | 38.153543, −101.411501 | Transcriptome (PRJNA513164) | No | No | No | – |
| *Pyxidicula operculata* | Organoconcha | Database Search | Brazil | –23.566339, −46.730632 | Transcriptome (PRJNA513164) | Yes* | No | No | – |
| *Planocarina carinata* | Hyalospheniformes | Database Search | France | 33.254160, −89.089728 | Transcriptome (PRJNA513164) | Yes* | No | No | – |
| *Nebela* cf. *tincta* | Hyalospheniformes | Database Search | USA | 33.254160, −89.089728 | Transcriptome (PRJNA513164) | Yes* | No | No | – |
| *Hyalosphenia elegans* | Hyalospheniformes | Database Search | France | 46.831247, 6.160270 | Transcriptome (PRJNA513164) | Yes* | No | No | – |
| *Hyalosphenia papilio* | Hyalospheniformes | Database Search | France | 46.831247, 6.160270 | Transcriptome (PRJNA513164) | Yes* | No | No | – |
| *Difflugia nodosa* Q7 | Longithecina | Environmental Isolate | The Netherlands | 52.281187, 5.118867 | Single-Cell Barcode COI (OQ631085) | Yes | No | No | – |
| *Lagenodifflugia vas* | Longithecina | Environmental Isolate | The Netherlands | 52.281187, 5.118867 | Single-Cell Barcode COI (OQ631084) | Yes | No | No | – |
| *Difflugia* sp. | Longithecina | Database Search | Brazil | –22.930797, −43.970812 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Difflugia compressa* | Longithecina | Database Search | Brazil | –23.565708, −46.728689 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Difflugia bryophila* | Longithecina | Database Search | USA | 33.254160, −89.089728 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Lesquereusia mimetica* | Longithecina | Database Search | Brazil | –22.692778, −43.030833 | Transcriptome (PRJNA513164) | Yes* | Yes* | No | – |
| *Argynnia* sp. Q6 | – | Environmental Isolate | Spain | 40.596694, −3.920694 | Single-Cell Barcode COI (OQ631083) | Yes | No | No | – |
| Unidentified Glutinoconcha M8 | Longithecina | Environmental Isolate | Spain | 40.750667, −4.065917 | Single-Cell Barcode COI (OQ631088) | Yes | No | No | – |
| Unidentified Glutinoconcha F24 | Longithecina | Environmental Isolate | Spain | 40.596694, −3.920694 | Single-Cell Barcode COI (OQ631087) | Yes | No | No | – |

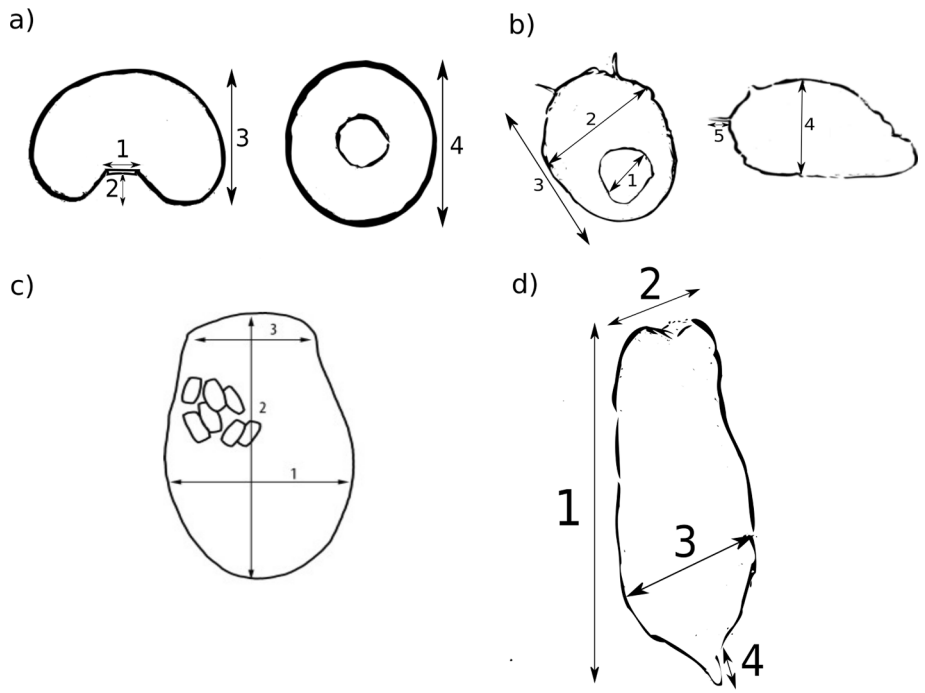


**Fig. 1.** Measurements taken for morphometrical analyses of the described species. (a) *Arcella uspiensis* – 1: aperture diameter, ad; 2: aperture height, ah; 3: test height, th; and 4: test diameter, td. (b) *Centropyxis blatta*. – 1: apertural diameter, ad; 2: test breadth, tb; 3: test length, tl; 4: test height, th; 5: spine length, sl. (c) *Heleopera steppica* – 1: test length, tl; 2: oral width, ow; 3: test length, tl. (d) *Cylindrifflugia periurbana* – 1: test length, tl; 2: test aperture, ta; 3: test width, tw; 4: Aboral spine, as.

minimum quality of 28 and length of 75 bp. We assembled using rnaSPAdes (http://bioinf.spbau.ru/en/spades_3_9). We included the assemblies in the database search for COI and SSU sequences described here.

- Database search

There are currently more than a hundred Arcellinida entries for transcriptomic data in NCBI. To increase the number of taxa in our phylogenetic reconstructions we mined data from taxa of interest in the NCBI database (excluding most Hyalospheniformes), using Blastn searches (Altschul et al., 1990) for each taxon from Sequence Read Archive (SRA) datasets using Arcellinida COI sequences as a query. We downloaded the first 100 results from each species and built contigs using Geneious Prime 2022 (https://www.geneious.com). We aligned the obtained sequences with other known COI sequences from Arcellinida, already available in the NCBI database, using MAFFT (Katoh et al., 2002) and visualised the results using SeaView (Gouy et al., 2010). We then built maximum likelihood phylogenies using IQ-TREE (Nguyen et al., 2015). We looked for COI sequences in assembled transcriptomes using the hmmsearch tool with an HMMER profile of Arcellinida-aligned sequences constructed by hmmbuild (Finn et al., 2011). Then, we analyzed the output and filtered out the contigs that did not correspond with the desired COI fragment. In the final analysis, we added sequences obtained through metabarcoding data obtained in (González-Miguéns, 2023) as well as sequences from characterised organisms (González-Miguéns et al., 2022a, 2022b).

All information about organisms that did not derive from this work (i.e. obtained through publicly available transcriptomic data) was obtained from the original publications. In that case, both optical microscopy and electron microscopy are indicated when available in the original source in Table 1 and Supplementary Table S2.

- Phylogenetic reconstruction

We aligned sequences using the MAFFT-LINSI setting, the most accurate from MAFFT method (Katoh et al., 2002). We trimmed the alignments using TrimAl (Capella-Gutiérrez et al., 2009) with a 30% gap threshold (gt). The reason behind the use of this low threshold was the shorter length of the environmental sequences. We searched for the best substitution model with Model Test, as implemented in IQ-TREE v. 1.6.12, and built a Maximum Likelihood tree with the same program using 1000 ultrafast bootstraps (Nguyen et al., 2015). We used four Lobosea sequences as an outgroup: *Saccamoeba* sp., LC102286, *Copromyxa protea*, LC102284, *Copromyxa cantabrigiensis*, LC102285, and *Copromyxa* sp., LC102283 (Kostka et al., 2017). We analyzed the effect of taxon sampling in our COI phylogenies by removing unreliable sequences with Guidance2 (guidance.tau.ac.il) with a default confidence score (of 0.6). Also, the eDNA sequences helped in stabilising the phylogeny, improving the support. For that reason, we kept the eDNA sequences in the phylogeny except for a few long branches (RA31_3, RA25, RA52, ASV_2388, ASV_2503). Finally, we built subtrees that represented each Arcellinida infraorder (except Hyalospheniformes, which have been detailed elsewhere; see Duckert et al., 2018) with the methodology described above. We also built an SSUrRNA tree for comparison with the literature. We added the newly produced sequences of *Heleopera steppica* sp. *nov.*, *Centropyxis blatta* sp. *nov.*, *Argynnia* sp., *Difflugia nodosa* Q7, *Lagenodifflugia vas* Q6 and Unidentified Glutinoconcha M8 and F24. We also looked for the SSU sequences through a blastn search against the Genbank database. We applied a 50% gt for trimming and used Lobosea sequences as outgroups: *Nolandella hibernica* JQ519510, two sequences of *Vermamoeba vermiformis* JQ519505 and AY502961, and *Echinamoeba thermarum* AJ489266 (Baumgartner et al., 2003, Lahr et al., 2011). The tree was constructed by maximum likelihood in IQ-TREE.

We also built subtrees (one for each suborder and, within Glutinoconcha, one for each infraorder) to remove the possible effects of different evolutionary paces in the different groups and to visualize more accurately the position of the new isolates; the numbers of isolates per tree are indicated in Table 2. All COI alignments had around 640 bp, except for the specific Excentrostoma alignment (360 bp, tree in Table 2), where most sequences derived from short environmental DNA

**Table 2**
Number of taxa included and size of the alignment for the complete phylogenies an each of the subtrees of the work. We made subtrees with different combinations of taxa from the Arcellinida lineages stabliished in Lahr et al., 2019 and González-Miguéns et al., 2022b.

| Infraorders | Taxa included | Alignment size (bp) | Conserved sites(%) | Parsimony informative sites (%) |
|---|---|---|---|---|
| **COI** | ~ | ~ | ~ | ~ |
| All taxa | 254 | 640 | 28 | 61 |
| Phryganellina, Organoconcha, Volnustoma | 24 | 640 | 47 | 36 |
| Excentrostoma | 108 | 360 | 42 | 48 |
| Cylindrothecina | 33 | 613 | 50 | 41 |
| Sphaerothecina | 81 | 632 | 44 | 49 |
| **SSU** | ~ | ~ | ~ | ~ |
| All taxa | 99 | 1100 | 24 | 62 |
| Phryganellina, Organoconcha, Volnustoma | 28 | 1865 | 40 | 49 |
| Excentrostoma | 19 | 1178 | 35 | 46 |
| Cylindrothecina | 19 | 1178 | 44 | 46 |
| Sphaerothecina | 42 | 1169 | 52 | 37 |

metabarcodes (340 bp). SSU sequence alignment sizes varied from 1100 (all taxa) to 1823 bp (Phryganellina + Organoconcha + Volnustoma). Both subtrees of Phryganellina + Organoconcha + Volnustoma recovered a topology that resembled the one obtained with phylotranscriptomic data (Lahr et al., 2019). However, the relationship between the species did not change from the total species alignment.

## 3. Results

- Microscopical observations

The new measurements of *Arcella uspiensis* sp.nov. (previously *Arcella intermedia,* n = 30) in culture are significantly smaller than the previous measurements obtained for the lineage (Porfírio-Sousa and Ribeiro et al., 2017), suggesting that organisms have shrunk after five years in culture (Fig. 2). This has been previously observed to happen with this culture in the laboratory (unpublished observation by D. Lahr). Both *Centropyxis blatta sp. nov.* (n = 30) and *Cylindrifflugia periurbana sp. nov.* (Previously *Cylindrifflugia acuminata,* n = 4) lineages are significantly smaller than the measures presented previously for similar morphotypes as *Centropyxis aculeata* (Lahr et al., 2008) and *Cylindrifflugia sp. cf. acuminata* (Ehrenberg, 1838) (Fig. 3, Supplementary Table S1). The dimensions of *Heleopera steppica sp. nov.* is consistent with the original description of *H. sylvatica* (Penard, 1890); however, it is smaller than the new specimens presented by Todorov & Bankov (2019) (Supplementary Table S1). From all the taxa included in this manuscript, nine were documented only with light microscopy and four with both light microscopy and scanning electron microscopy (SEM).

- Molecular work and database search

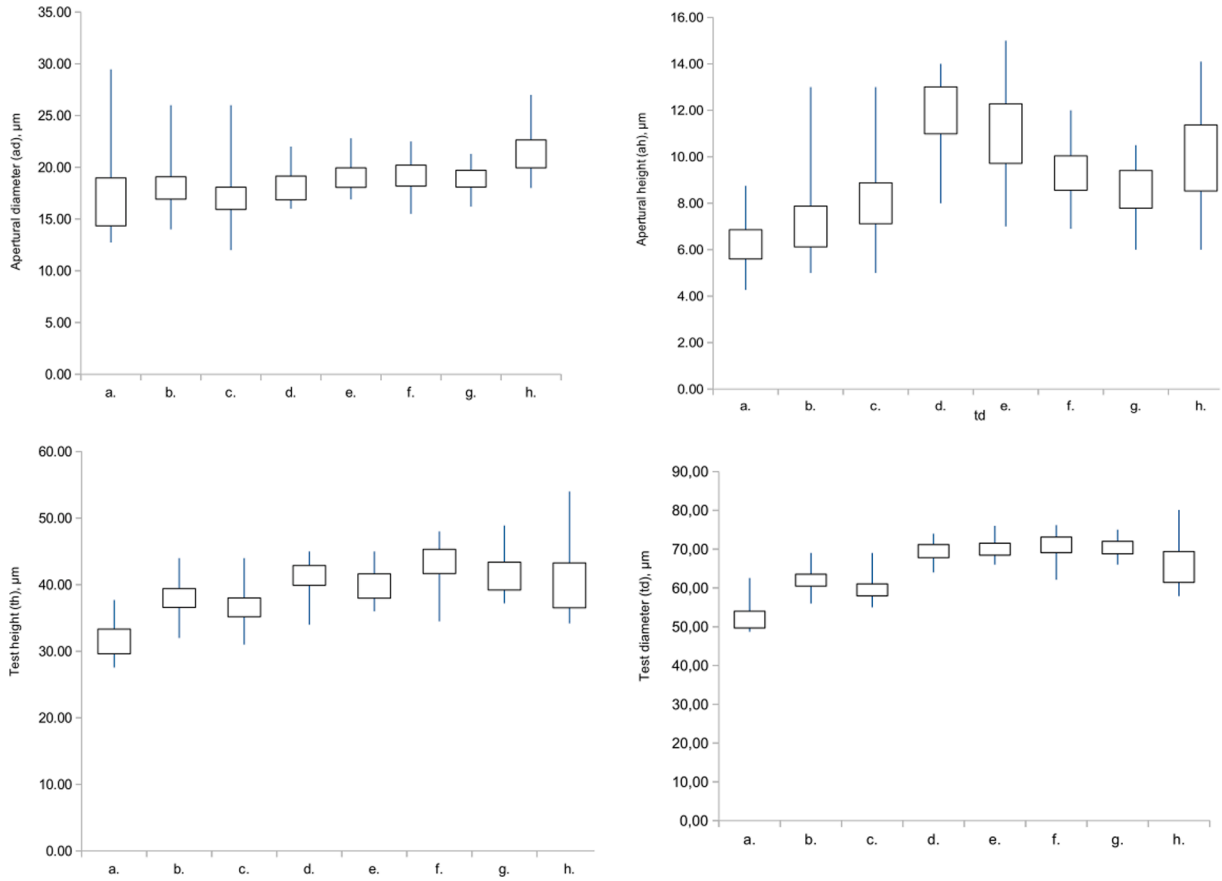


**Fig. 2.** Measurements of *Arcella uspiensis* type and comparison with previously described "*Arcella intermedia*" morphotypes. We took pictures from the lateral and apertural view, and measured the diameter and height of the test (td and th), also the diameter and height of the aperture (ad and ah). The original description measurements provided by Deflandre of *A. hemisphaerica intermedia* (1928) are on average smaller than the species here (diameter 48–68 m). a. *Arcella uspiensis* current culture; b. "*Arcella intermedia laevis*" culture (Porfírio-Sousa and Ribeiro et al., 2017); c. "*Arcella intermedia*" culture (Porfírio-Sousa and Ribeiro et al., 2017); d. Arcella intermedia"wild type" P1 (Porfírio-Sousa and Ribeiro et al., 2017); e. "*Arcella intermedia*" wild type P2 (Porfírio-Sousa and Ribeiro et al., 2017); f. "*Arcella intermedia*" P1 (Tsyganov and Mazei, 2006); g. "*Arcella intermedia*" P2 (Tsyganov and Mazei, 2006); h. "*Arcella intermedia*" P3 (Tsyganov and Mazei, 2006).

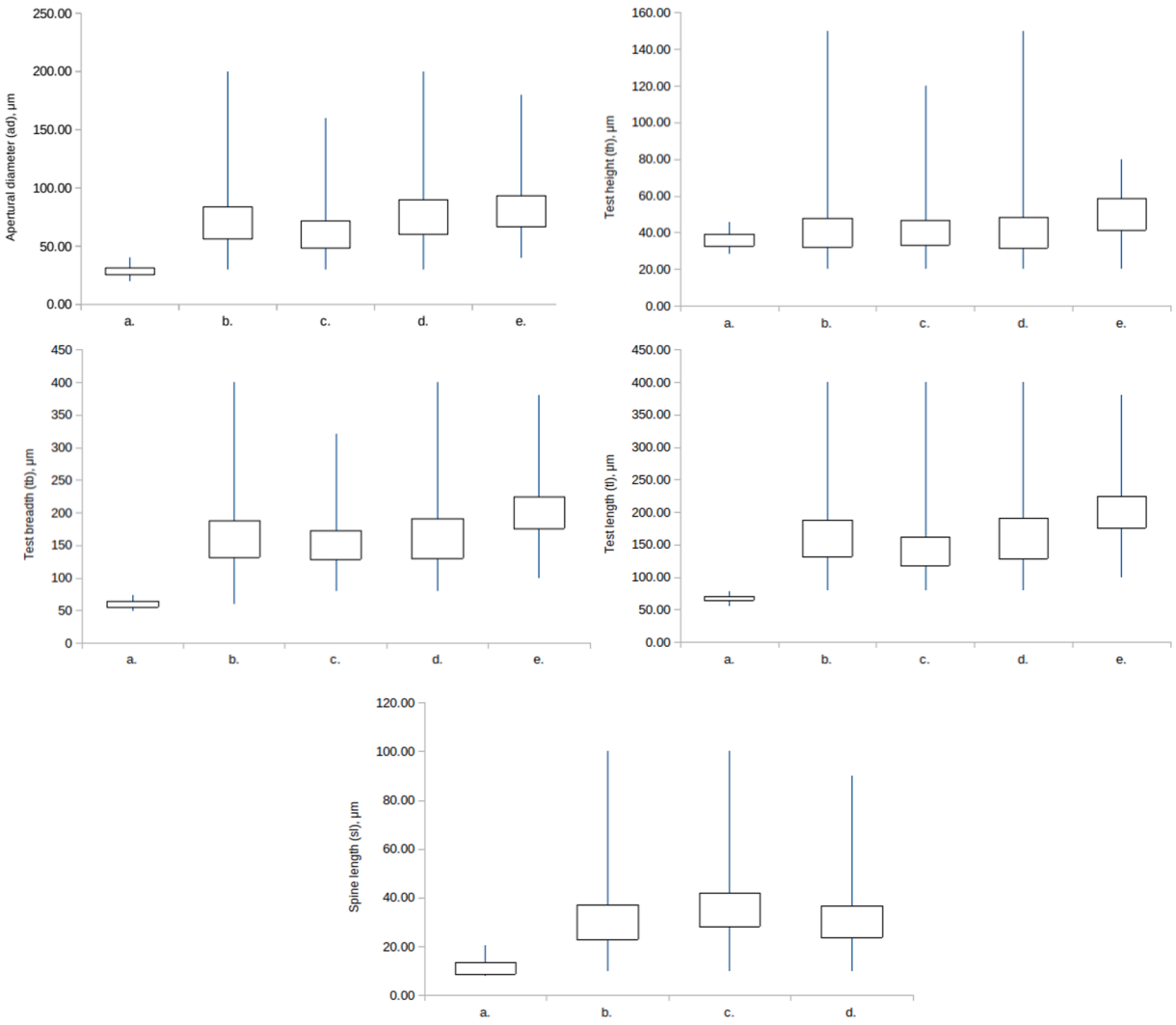


**Fig. 3.** Morphometric measurements from *Centropyxis blatta* in comparison with previously described *Centropyxis* morphotypes (Lahr et al., 2008). We took pictures from ventral and lateral view, and we measured the aperture diameter (ad), test height (th), test breadth (tb), test length (tl), spine length (sl). a. *Centropyxis blatta*; b. *Centropyxis* - All taxa (Lahr et al., 2008); c. *Centropyxis aculeata* (Lahr et al., 2008); d. *Centropyxis discoides* (Lahr et al., 2008).

We obtained 10 new single-cell PCR sequences of Arcellinida, 9 from COI and one for the ribosomal gene (SSUrRNA) (Table 1). Amongst COI sequences, 7 were obtained directly with the HCO-LCO (641 bp) and 2 with the semi-nested approach ArCOIR-LCO (359 bp). Also, 2 sequences were sorted from transcriptomic data sequenced data of *Phryganella* specimens (PRJNA945922, Table 1). Moreover, we obtained one partial sequence for SSU by PCR amplification, and 4 from transcriptomic data (PRJNA945922, Table 1)All the barcoded single cells are indicated in Table 1. We explored publicly available sequences SRA data from 37 different lineages and obtained COI sequences from 12 and SSU from 18 of them (Supplemntary Table S2). The SRA data analysed contain all information about the original bioproject; sample description and sequencing platform used are in Supplementary Table S2. Moreover, we added a total of 139 eDNA sequences obtained in (González-Miguéns, 2023).

- Phylogenetic reconstruction

The trees resulting from phylogenetic reconstructions (both COI and SSU) are illustrated in Fig. 4. In the end, we had a total of 254 sequences for the final COI alignment, with 640 base pairs after trimming (-gt 0.3). We based our SSU phylogeny on the one present in Lahr et al., 2019 but added the newly acquired sequences of *Arcella uspiensis sp. nov.*, *Centropyxis blatta sp. nov.*, *Heleopera steppica sp. nov.*, two *Phryganella paradoxa*, and one *Phryganella acropodia* sequences. The final SSUrRNA alignment had 99 sequences with 1100 base pairs after trimming (-gt 0.5) (see Figs. 5-9).

Both COI and SSU phylogenies showed weak support for the deep relations between Arcellinids sub- and infraorders (Fig. 4). Nevertheless, the SSU tree placed Volnustoma as a sister group to all other Glutinoconcha, with a good support (99%), as in a previous phylogenomic work (Lahr et al., 2019). COI was better suited for fine level differentiation of closely-related lineages (i.e. species level); for instance, *Heleopera sylvatica* PRJNA513164 and *Heleopera steppica* sp. *nov.* had a sequence divergence of 1,15% with SSU versus 11,44% with COI. All the other Arcellinida infraorders were recovered in both phylogenies, with moderate to high support (minimum bootstrap support 75%), and are indicated by colour boxes with the name of the group in Fig. 4.

In this work, we obtained for the first time COI sequences for suborders Phryganellina, Organoconcha, and infraorder Volnustoma. All the Phryganellina sequences were grouped robustly (bootstrap = 99%). The SSU phylogeny also recovered the suborder Phryganellina with full support and also efficiently separates genus *Cryptodifflugia* and *Phryganella*. It seems, however, that *Phryganella paradoxa* appears paraphyletic and includes probably several species. The node at the base of infraorder Organoconcha appears paraphyletic as both *Spumochlamys* sequences

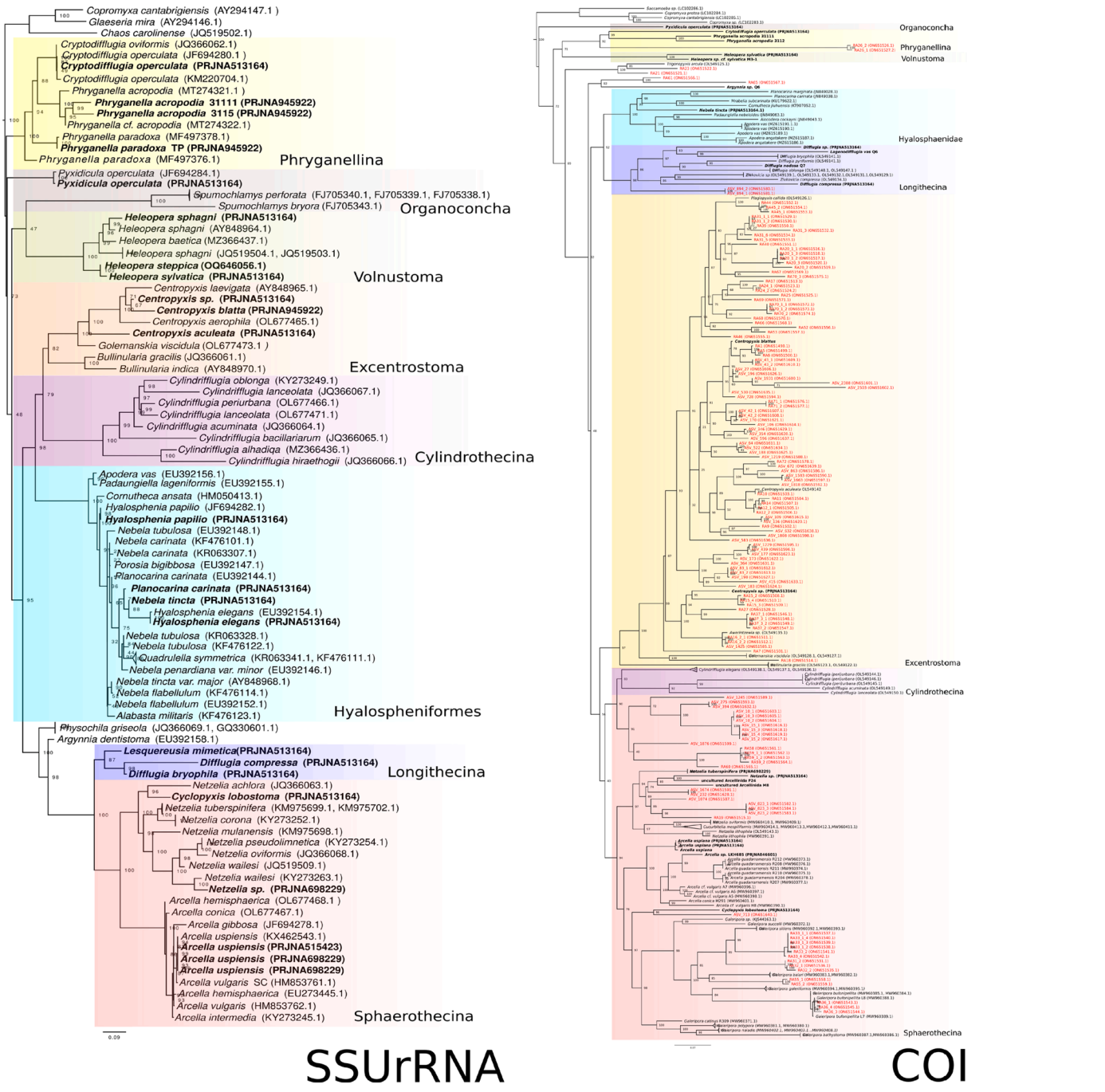


**Fig. 4.** Arcellinida phylogenetic reconstructions of SSUrRNA (left) and COI (right) markers. The final SSUrRNA alignment had 99 sequences with 1100 base pairs after trimming (-gt 0.5). The COI alignment had a total of 254 sequences with 640 base pairs after trimming (-gt 0.3). Both trees are rooted with Amoebozoa taxa. The new sequences from both transcriptomic and sampling sources shown in bold. The major lineages of Arcellinida, established in Lahr et al., 2019 and González-Miguéns, 2022b are indicated in the colour boxes.

branch with Volnustoma in the SSU rRNA tree. This result is not supported by multigene trees, and is most likely an artefact that results from an attraction between strongly divergent SSU sequences of *Spumochlamys* and *Heleopera*.

Agglutinating species without any special extra feature were traditionally placed within genus *Difflugia*; however, recent phylogenetic work demonstrated that the genus was not monophyletic, and that species were distributed in all Glutinoconcha infraorders (González Miguéns et al., 2022b). Here, sequences from both *Difflugia* species obtained from the transcriptomes in Lahr et al., 2019 (*Difflugia compressa* and *D.* sp.) branch within Longithecina together with the type species of the genus, *D. pyriformis*. *Cylindrifflugia periurbana*, which resembles a smaller version of *C. acuminata* branches robustly with the latter. The still phylogenetically uncharacterized *Lagenodifflugia vas* belonged in Longithecina, and branched together with *D. bryophila,* which shares a similar neck morphology. However, certain isolates did not branch where they would be expected based on morphology. *Centropyxis blatta* did not appear to be closely related to *C. aculeata* in the COI tree. Both COI sequences of unidentified Glutinoconcha F24 and M8 branch robustly within the Sphaerothecina, despite their shell outline which strongly reminds of some species of genus *Difflugia* (Longithecina). Finally, the position of *Argynnia* sp. was still not characterised, but

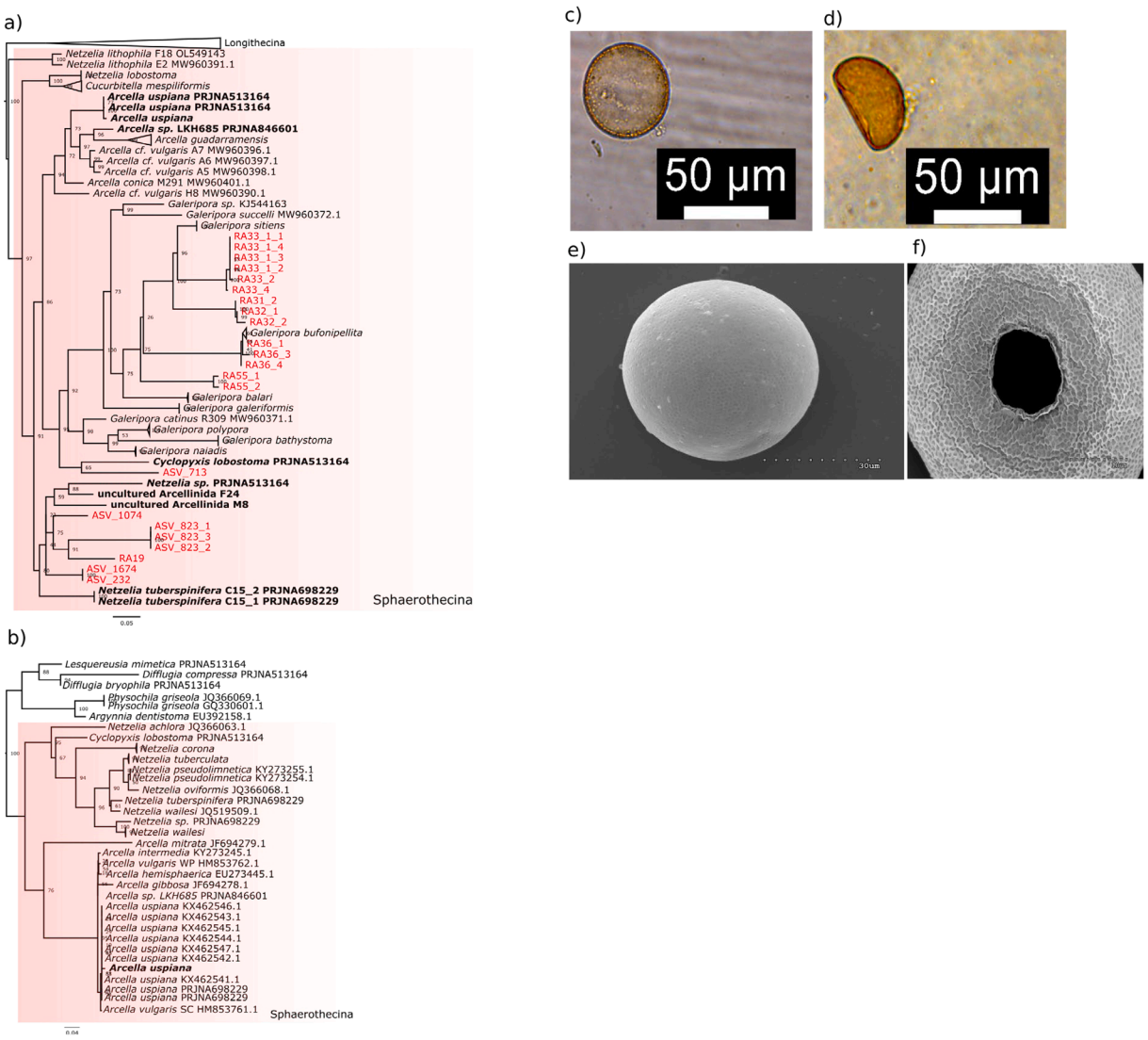


**Fig. 5.** Sphaerothecina subtrees and phylogenetic placement of *Arcella uspiensis* (a) COI subtree of Sphaerothecina taxa with Longithecina as external group. (b) SSU subtree of Sphaerothecina with Longithecina as external group. (c) Light microscopy of *A. uspiensis* in apertural view (d) Light microscopy of *A. uspiensis* in lateral view. (e) Electron microscopy of aboral view of *Arcella uspiensis*. (f) Electron microscopy of oral view of *Arcella uspiensis*.

branched together with two environmental clones (RA61 and RA 65).

Phylogenetic analysis of Arcellinida always improve species discrimination. For example, in both phylogenies there are two distinct *Cylindrifflugia acuminata* groups of sequences, one isolated in Spain and another from Bulgaria. In SSU, *Cylindrifflugia acuminata* (OL677466) from Spain falls separated from *Cylindrifflugia acuminata* (JQ366064) from Bulgaria. In COI, *Cylindrifflugia acuminata* (OL549144, OL549145, OL549146) from Spain are also separated from *Cylindrifflugia acuminata* (OL549149) from Bulgaria. COI marker also improve species-level discrimination with respect to SSU, as shown for instance in the Sphaerothecina *Arcella conica* (MW960401) and *Arcella* cf. *vulgaris* (MW960390). However, the relationships at the infraorder level are lost or less supported with COI. For example, the node at the base of the family Netzeliidae is lost in the Sphaerothecina subtree. Finally, for Excentrostoma, most of the supports did not change much between the original and the subtree, but a few deeper nodes, for example the one that associates *Centropyxis blatta* with *Centropyxis aculeata,* became better supported. All the subtrees of SSUrRNA also had similar sizes after trimming with more than a thousand base pairs (Table 2).

## 4. Taxonomic actions

We proceeded to describe as new species either (1) organisms whose morphology did not match with previous descriptions or (2) organisms whose original description was ambiguous and did not allow finding back the original species. We only described taxa for which we obtained sequences from both SSU and COI markers, and a micrograph of sufficient quality allowing the identification of the original traits (light or SEM, scanning electron microscopy). We are describing four new species of Arcellinida; *Arcella uspiensis sp. nov.*, *Centropyxis blatta sp. nov.*, *Heleopera steppica sp. nov.* and *Cylindrifflugia periurbana sp. nov.*.

Descriptions here follow ICZN usage as is traditional in testate amoebae, but attempt to reconcile with the now more commonly used Phylocode for protists. The testate amoeba community has been discussing this issue and justification for these actions can be found elsewhere (Lahr et al. 2012, 2019).

Amorphea Adl et al. 2012
Amoebozoa Luhe 1903, *sensu* Cavalier-Smith 1998
• Tubulinea Smirnov et al. 2005
•• Elardia Kang et al. 2017
••• Order Arcellinida Kent 1880
•••• Suborder Glutinoconcha Lahr et al 2019

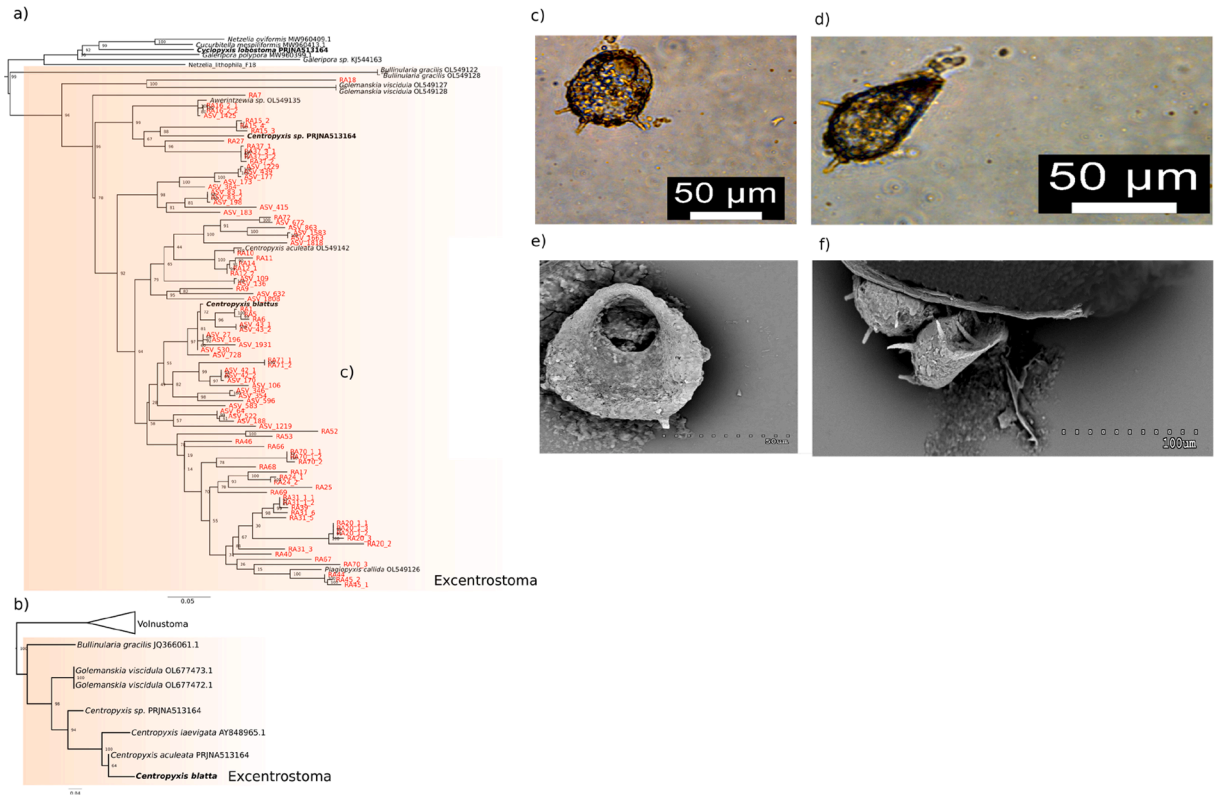


**Fig. 6.** Excentrostoma subtrees and evidence of the new species in the group *Centropyxis blatta* (a) COI subtree of Excentrostoma taxa with Other Arcellinida sequences as outgroup. (b) SSU subtree of Excentrostoma with Cylindrothecina as outgroup. (c) Light microscopy of apertural view of *Centropyxis blatta.* (d) Light microscopy of lateral view of *Centropyxis blatta* (e) Electron microscopy of apertural view of *Centropyxis blatta*. (f) Electron microscopy of lateral view of *Centropyxis blatta*.

●●●●● Infraorder Sphaerothecina Kosakyan, Lara and Lahr 2016
●●●●●● Family Arcellidae Ehrenberg 1843
●●●●●●● Genus *Arcella* Ehrenberg 1832

***Arcella uspiensis* n. sp. Giulia M. Ribeiro, Alfredo L. Porfírio-Sousa, Daniel J.G. Lahr, Enrique Lara**

Included organisms previously studied:
*Arcella* cf. *intermedia* PRJNA513164

**Zoobank registration**: F5D2DD9E-ECDA-4021–9459-FD723925F1DA

**Holotype**: deposited at the Zoology Museum of the University of Sao Paulo - add the code in the revision

**Specific diagnosis**: test diameter: 64–76 µm, average 69.85 µm (N = 52); test height 34–45 µm, average 40.75 µm (N = 52); aperture diameter 16–22.8 µm, average 19.2 µm (N = 52). aperture height: 7–15 µm, average 11.45 µm (N = 52).

General shell shape hemispherical, with a circular shape and composed of self-secreted organic material. Shell composed of building units consisting of hexagonal alveoli with pores on the sides. Aperture is invaginated, surrounded by an organic collar, located in a centric position and its diameter is about half of shell width. 2 nuclei. Variable number lobose-shaped pseudopods.

**Diagnosis with closely related species**: *Arcella uspiensis* can be diagnosed by its specific COI and SSU rRNA sequences and by its phylogenetic placement. Measurements of individuals directly isolated from the environment are concordant with *Arcella intermedia* as defined in Tsyganov and Mazei (2006). However, as argued by these authors, *A. intermedia* includes a wide number of forms whose taxonomic position is still unclear. The measurements provided by Deflandre in his original description as *A. hemisphaerica intermedia* (1928) are on average smaller than the species here (diameter 48–68 µm), although Deflandre did not provide single cell measurements that can be used to perform a morphometric analysis. In addition, he cites as *terra typica* the Jura Mountains in France, very far away from the location where *A. uspiensis* has been first isolated. For these reasons, we decided to erect *A. uspiensis* as a species morphologically resembling but different from the original *A. intermedia.*

**Habitat**: Freshwater sediment, collected in an artificial pond in an Atlantic rainforest fragment within the campus of the University of São Paulo.

**Type locality**: Brazil, State of São Paulo, São Paulo City (23°33′ S, 46°43′ W).

**Etymology**: The species epithet is derived from a latinized form of the abbreviation of the University of São Paulo (USP), where the types have been isolated.

**Specific remarks:** While shell dimensions are well conserved in specimens directly isolated from natural samples, shell size tends to decrease after successive subculturing steps (mostly after one year – Fig. 2).

●●●●● Infraorder Excentrostoma Lahr et al. 2019.
●●●●●● Family Centropyxidae Deflandre, 1953
●●●●●●● Genus *Centropyxis* Stein, 1857

***Centropyxis blatta n. sp.* Giulia M. Ribeiro, Enrique Lara**

**Zoobank registration**: E8412E75-A11F-4979-AFE4-E83D2D808823

**Holotype**: MA-amoeba11272

**Specific diagnosis:** shell length: 54.77–77.31 µm, average 66.34 µm (N = 25); shell width 28.20–45.70 µm, average 36.41 µm (N = 25); aperture 19.98–40.42 µm, average 27.59 µm (N = 25). Shell breadth: 48.43–74.03, average 59.99 (N = 25). Spine length: 7.66–20.39 µm, average 11.65 µm (N = 19).

Shell yellow, brown or colourless, becomes darker with ageing. Ovoid in apertural view. Aperture oval and sub-terminal. Rounded

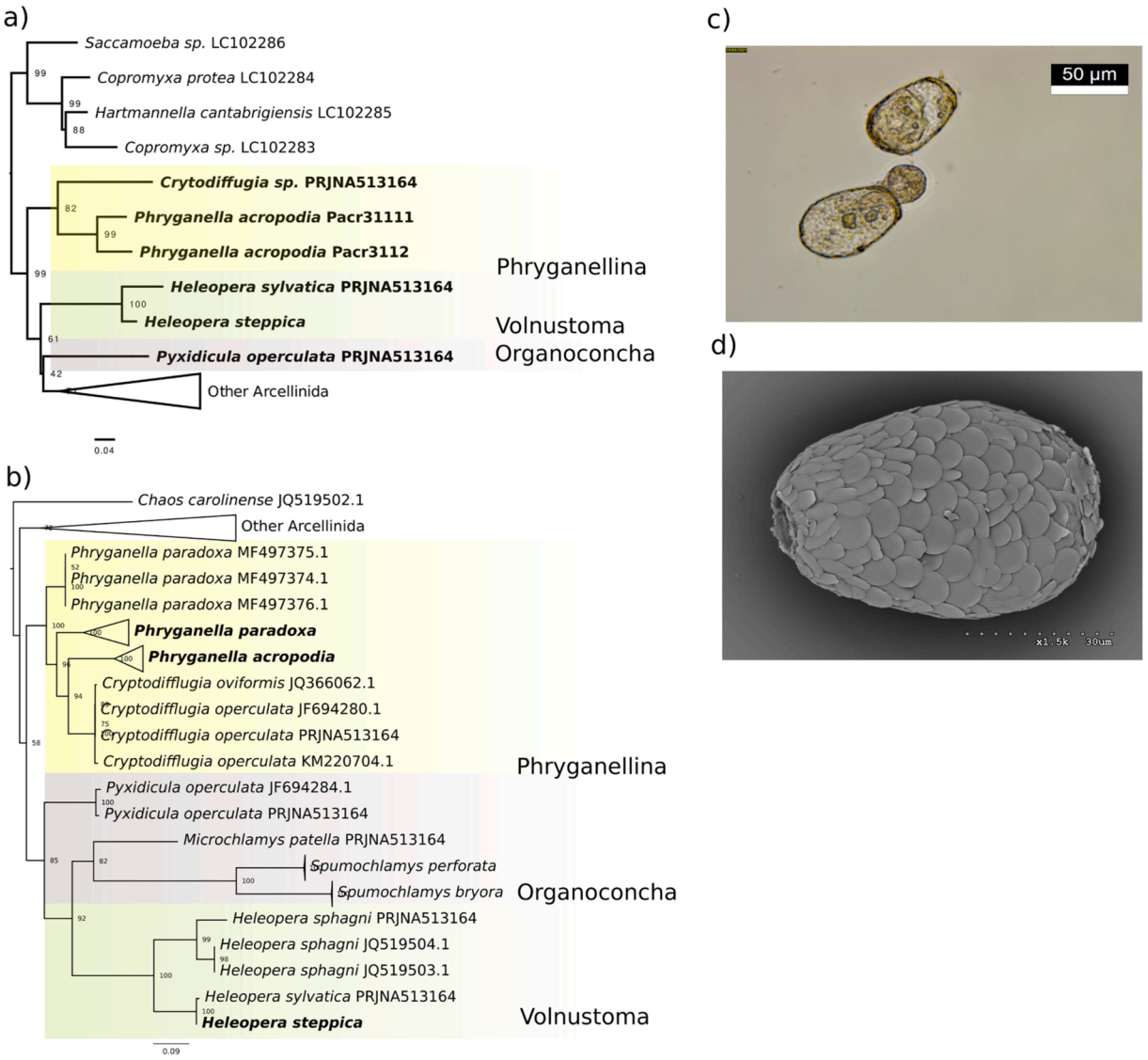


**Fig. 7.** Phryganellina + Volnustoma + Organoconcha subtrees and evidences of the new species in the group *Heleopera steppica* (a) COI subtree of Phryganellina + Volnustoma + Organoconcha taxa with Other Amoebozoa sequences as external group. (b) SSU subtree of Phryganellina + Volnustoma + Organoconcha with Other Amoebozoa as outgroup. (c) Optical microscopy of *Heleopera steppica* specimens (d) Electron microscopy of lateral view of *Heleopera steppica*.

fundus (=opposed to the aperture) and flattened near the aperture in lateral view. Variable number of spines (0–4, n = 25). Shell composed of exogenous material (usually mineral particles) attached with an organic cement.

**Diagnosis with closely related species**: The general morphology of *Centropyxis blatta* fits within the *Centropyxis aculeata* broad morphotype (Amesbury et al., 2018). However, it is considerably smaller than the type description by Ehrenberg (1838), who cites a length of 1/18 Linie (118 µm). The general shell outline of *Centropyxis aculeata oblonga* resembles *C. blatta* but is twice as large (Deflandre, 1929). At the current state of knowledge of Excentrostoma diversity, there is no closely resembling species described. Additionally, *Centropyxis blatta* can be diagnosed by its specific COI and SSU rRNA sequences and by its phylogenetic placement.

**Habitat**: Freshwater sediment, collected in the eutrophic pond of an urban park.

**Type locality**: Spain, Madrid, Parla (40°14′ N, 3°46′ W), in a pond in the "Parque del Universo".

**Etymology:**

Blatta means cockroach in Latin, as a reference to the small size, dark coloured and horned test, as well as the fact that it lives in nutrient rich environments, like the insects.

●●●●● Infraorder Volnustoma Lahr et al. 2019.
●●●●●● Family Heleoperidae Jung 1942
●●●●●●● Genus *Heleopera* Leidy, 1879
***Heleopera steppica* Useros and Lara**
Included previously studied organism *Heleopera* sp. cf. *sylvatica* M3-1

**Zoobank registration**: 0992BDAB-59CB-4B01-AD68-73A32DA8BF45

**Holotype**: MA-amoeba11271

**Specific diagnosis**: shell length: 60.7 – 71.0 µm, average 66.05, standard deviation (SD) 2.49 µm (N = 24); Maximum width 27.9 – 46.1 µm, average 41.73, SD 3.49 µm (N = 24); Oral width 20.3 – 32.7 µm, average 24.91, SD 3.03 µm (N = 24); Distance from point of maximum with to bottom of the test 15.6 – 26.0 µm, average 20.93, SD 2.95 µm (N = 24).

Transparent test, colourless or yellowish. Ovoid shape, but vertically compressed and with a flatter oral side. Narrow and elliptical terminal oral aperture with a thin cement collar. Test composed of circular and elongated siliceous shell plates of various sizes, obtained from diatom

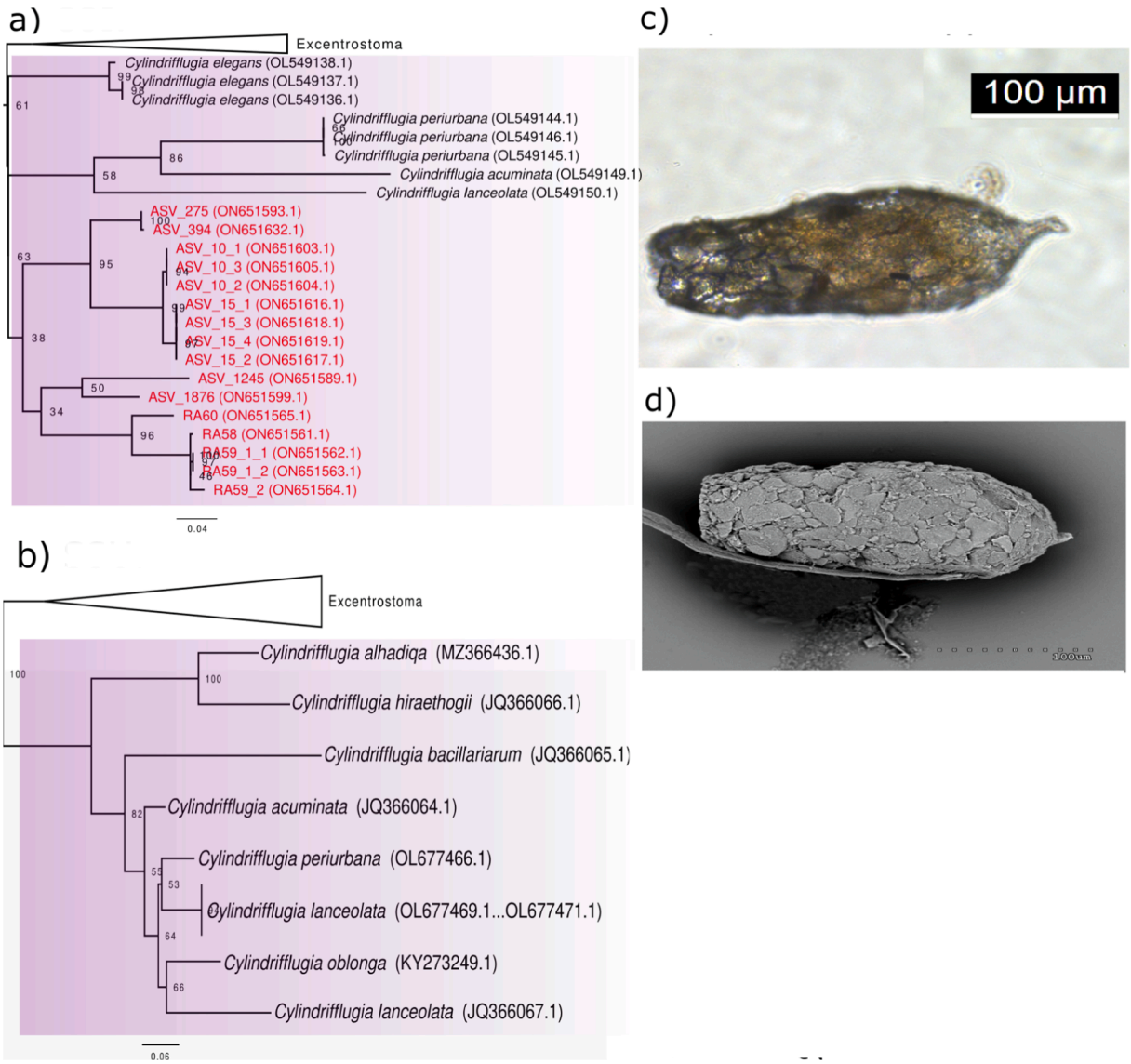


**Fig. 8.** Cylindrothecina subtrees and evidences of the new species in the group *Cylindrifflugia periurbana* (a) COI subtree of Cylindrothecina taxa with Other Arcellinida sequences as external group. (b) SSU subtree of Cylindrothecina with Excentrostoma + Hyalospheniformes as external group. (c) Optical microscopy of lateral view of *Cylindrifflugia periurbana* (d) Electron microscopy of lateral views of *Cylindrifflugia periurbana*.

frustules and Euglyphida shell scales, on which they prey.

**Diagnosis with closely related species**: *Heleopera steppica* can be diagnosed by its specific mtDNA (COI) and nuclear (SSU rRNA) sequences and by its phylogenetic placement. It is morphologically very similar to *H. sylvatica* and falls within the size range given in Penard (1890) original description of 50 – 75 µm, although it is smaller than the range given in Todorov & Bankov (2009), 88 – 108 µm. The original description of Penard illustrated individuals with perpendicularly attached siliceous plates on the margin of the aboral region, which are absent in *Heleopera steppica.*

The *terra typica* of *H. sylvatica* Penard 1890 is Wiesbaden, Germany, where it was collected in temperate forest mosses. *H. steppica* has been found in a completely different environment, mosses growing on saline gypsum soils in shrublands from central Spain under a Mediterranean-continental, semi-arid climate. The difference between habitats also suggests that *H. steppica* and *H. sylvatica* are different species.

**Habitat:** Found in mosses from Mediterranean-continental shrublands and grasslands growing on gypsum soils.

**Type locality:** Salobral de Ocaña, Spain (39°59′30″N, 3°36′55″W)

**Etymology:** The name derives from steppe, making reference to the type of environment in which it was found.

••••• Infraorder Cylindrothecina González-Miguéns et al., 2022.

•••••• Family Cylindrifflugidae González-Miguéns et al., 2022

••••••• Genus *Cylindrifflugia* González-Miguéns et al., 2022

***Cylindrifflugia periurbana n. sp.*** **Rubén Gonzáles-Miguéns, Giulia M. Ribeiro, Enrique Lara**

Included organisms previously studied:

*Cylindrifflugia acuminata* voucher P1, *Cylindrifflugia acuminata* voucher P4, *Cylindrifflugia acuminata* voucher P5.

**Zoobank registration**: B6D52AE1-5B64-4540-A7AF-B27CC362A3DE

**Holotype**: MA-amoeba11274

**Specific diagnosis**: shell length: 249.8–264.6 µm, average 255.25 µm (N = 4); shell width 104.3–118.9 µm, average 112.9 µm (N = 4); aperture 55–61.6 µm, average 59.7 µm (N = 4). Aboral spine: 26.41 µm (N = 4)

General shell shape is cylindrical-pyriform, with a circular section and composed of exogenous material (usually sand grains) attached with an organic cement. Aperture diameter equal to or smaller than the width

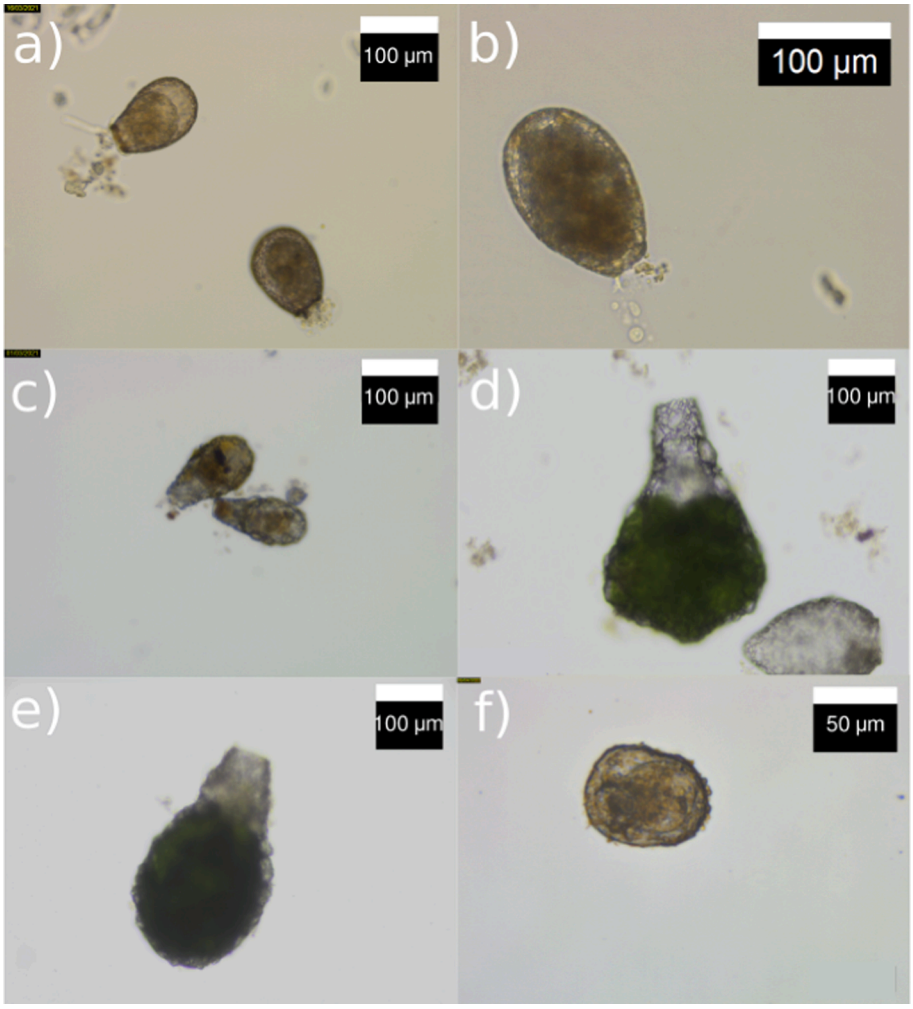


**Fig. 9.** Pictures of the new Arcellinid specimens that have been barcoded with COI and whose sequences have been added to the database. (a) *Argynnia* sp. (b) *Argynnia* sp. (c) Unidentified Glutinoconcha M8 (d) *Difflugia nodosa* (e) *Lagenodifflugia vas* (f) *Centropyxis blatta*. Magnification 400X. Scale bars are 50–100 um.

of the shell, but not larger. Rounded fundus, with a spine.

**Diagnosis with closely related species**: *Cylindrifflugia periurbana* can be diagnosed by its specific mtDNA (COI) and nuclear (SSU rRNA) sequences and by its phylogenetic placement. Differs morphologically from *Cylindrifflugia lanceolata* from its overall pyriform shell shape (thinner next to the aperture and wider fundus) and the presence of a spine. It is also considerably smaller than *Cylindrifflugia acuminata* (Ehrenberg, 1838), for which lengths of 333 µm (1/3 mm) and 375 µm (1/6 Linie) have been reported.

**Habitat**: Freshwater sediment, collected in the eutrophic pond of a urban park.

**Type locality**: Spain, Madrid, Parla (40°14′ N, 3°46′ W). Pond in the Parque del Universo.

**Etymology**: The name is derived from the Greek peri (around) and Latin urbs ("city") + -āna, "belonging to a city". We propose this name because of the type locality of this species, an urban pond.

## 5. Discussion

- Arcellinida species descriptions

In this work, we were also able to collect data for the description of four new Arcellinida species that occupy strategic positions in the Arcellinida phylogeny (*Arcella uspiensis sp. nov., Centropyxis blatta sp. nov., Cylindriflugia periurbana sp. nov., and Heleopera steppica sp. nov.*).

The morphology of *Arcella uspiensis* has a typical shape for the genus *Arcella* with a hemispheric, flattened on the oral side shell. Likewise, the morphotypes *A. intermedia, A. hemisphaerica, A. rotundata, and A. gibbosa* with all their described varieties and the potential phenotypic plasticity form a complex of forms that can only be resolved using a combination of molecular and morphometric approaches (Tsyganov and Mazei, 2007). Because original descriptions are not accurate enough to allow finding back the species in the environment, it is necessary in this case to either process lectotypifications or to new species description. The first option only makes sense when there is a reasonable chance that the new organisms isolated are conspecific with the first descriptions. Otherwise, and this is the case for *A. uspiensis* which has been found in a completely different ecosystem and localities than the resembling *A. intermedia*, new species need to be described. Such a process is indispensable for the reproducibility of manipulative experiments and the conclusions drawn from ecological observations.

In COI reconstruction, *Centropyxis blatta* falls together with a group of environmental sequences derived from freshwater sediment (González-Miguens et al., 2023); it appears only far related from the *C. aculeata* sequence from Gonzalez Miguéns et al., 2022 (OL549142). We found two distinct species that fit within the *Heleopera sylvatica* morphotype but which are genetically well distinct. Data for this morphotype come from the newly described species (*Heleopera steppica*) as well as from a transcriptome obtained from a Brazilian isolate (PRJNA513164). Another small-sized, genetically distinct *Heleopera* has been described recently, *H. baetica* (Soler-Zamora et al., 2021) from cyanobacterial mats in a karstic cave in Southern Spain. The diversity of

this genus of morphologically similar species has also been most certainly overlooked.

One new *Difflugia*-like sequence (*Cylindrifflugia periurbana*) branched within the Cylindrothecina, forming a robust clade with *C. lanceolata, C. oblonga* and *C. acuminata*, to which it resembles superficially; the general elongated shell as well as the presence of a horn on the fundus seems to remain relatively constant through evolution, in contrast to the examples presented above.

- New lineages in Arcellinida COI Database

In this work, we intended to expand the molecular database of Arcellinida towards groups that remain under-characterised to facilitate the interpretation of the phylogenetic tree. This enlarged database will also help interpreting environmental sequences obtained in metabarcoding studies. Currently, COI has become the most widely applied genetic marker in Arcellinida. Its taxonomic accuracy and its relative ease of amplification make it a perfectly suited marker for application to metabarcoding studies (González-Miguéns et al., 2023). In total, we added 23 COI sequences taken either from barcoded single cells or from transcriptomic data. We present here, for the first time, sequences from the suborders Phryganellina and Organoconcha, and the infraorder Volnustoma (Lahr et al., 2019). We also provided new COI sequences for all Arcellinid groups, emphasizing underexplored Glutinoconcha infraorders. These new data, accompanied by illustrations and sampling localities (when possible, measurements), will considerably help to place back sequences from ASVs that could not be assigned to any group for now (González-Miguéns et al., 2023).

In addition, we added ten sequences from organisms found in the Neotropics, a biogeographical realm that has been undersampled but holds great potential for new species discovery (Burdman et al., 2021). Most testate amoebae research, involving molecular work was conducted in the Northern hemisphere, and was associated with *Sphagnum* mosses in peatlands (Lara et al., 2008; Ruggiero et al., 2020; Singer et al., 2018; 2019). Our sequences from Neotropical ecosystems will constitute the first steps for a more extensive investigation of South American Arcellinida diversity.

Enriching infraorders with new sequences helps understanding the morphological evolution of the organisms and reconstruct the evolutionary history of the whole order Arcellinida. In contrast with previous expectations, test shape is not systematically conserved within deep lineages but varies in line with environmental pressure (González-Miguéns et al., 2022a). For instance, two morphotypes described here as "unidentified Glutinoconcha" (M8 and F24) were phylogenetically assigned to Sphaerothecina based on their COI sequence even though their shells did not have the round morphology supposed to characterise that infraorder (Kosakyan et al., 2016). The fact that shell traits vary along with the environment will play a most important role in inferring niche transitions in order Arcellinida as a whole.

The relationship between niche and shell outline is pervasive along the Arcellinida tree and responsible for evolutionary convergences between species, or genera, at all taxonomic levels making genus *Difflugia* polyphyletic (González-Miguéns et al., 2022b). Genus *Difflugia*, as it was understood before the first molecular studies (Gomaa et al., 2012, Gomaa et al., 2015) included a wide number of species with mineral agglutinated shell and a variable morphology and ecology. Successive works dismantled this genus, which included species in almost all Glutinoconcha infraorders except (as far as we know) Volnustoma (González-Miguéns et al., 2022b). In this work, we increased the number of *Difflugia*-like documented taxa mainly for the COI marker. The new sequences enriched infraorder Longithecina with data deriving from the phylogenomic study by Lahr and collaborators (2019), an unidentified species *Difflugia sp* and *Difflugia* cf. *compressa*. The species *Difflugia nodosa* and *Lagenodifflugia vas* derive from single cell isolates and branch also within this infraorder. Then, both species that host endosymbiotic algae, *Difflugia nodosa*, and *D. pyriformis* branch together Given the complex host-symbiont metabolic relationships, such association can be expected to occur only a limited number of times in the evolutionary history of the groups (Gomaa et al., 2014).

This lack of phylogenetic conservatism in the morphological patterns in Arcellinida may lead to false interpretations of diversity when only morphology is taken into account. Environmental studies showed that Excentrostoma is extremely diverse and probably composed mostly of species that could be assigned to genus *Centropyxis.* Ecological studies show that several species such as *Centropyxis aerophila* or *C. aculeata* are almost ubiquitous in terrestrial and aquatic environments, respectively (Baessolo, 2012, Roe et al., 2010, Souto et al., 2021). It is therefore most likely that these ubiquitous broad morphotypes are composed of many species. A similar case might occur with species from the genus *Heleopera*, where several morphologically similar species occupy different niches and are genetically distinct.

## Data availability

All alignments and phylogenies are shared on github (https://github.com/GiuliaRibeiro/ArcellinidaDatabase)

## Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. Funding for this work was also provided by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) fellowships to DJGL (#2013/04585); GMR (#2019/22109-0, #2020/15705-3); APS (#2019/22692-8, #2021/09529-0). KD was funded by the German Research Foundation grant 399699069 - DU 1863/1-1. This study was funded by the Spanish Government through the projects PGC2018-094660B-I00 and PID2021-128499NB-I00 10.13039/501100011033/, (MCIU/AEI/FEDER,UE).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejop.2023.126013.

## References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. JMB 215 (3), 403–410.

Amesbury, M.J., Booth, R.K., Roland, T.P., Bunbury, J., Clifford, M.J., Charman, D.J., Elliot, S., Finkelstein, S., Garneau, M., Hughes, P.D., et al., 2018. Towards a holarctic synthesis of peatland testate amoeba ecology: Development of a new continental-scale palaeohydrological transfer function for North America and comparison to European data. Quat. Sci. Rev. 201, 483–500.

Arrieira, R.L., Alves, G.M., Schwind, L.T., Lansac-Toha, F.A., 2015. Local factors affecting the testate amoeba community (Protozoa: Arcellinida; Euglyphida) in a Neotropical floodplain. J. Limnol. 74 (3).

Atasiei, D., Nasser, N.A., Patterson, C.W., Wen, A., Patterson, R.T., Galloway, J.M., Roe, H.M., 2022. Impact of post-tropical storm Arthur (2014) on benthic Arcellinida assemblage dynamics in Harvey lake, New Brunswick, Canada. Hydrobiologia 849 (13), 3041–3059.

Baessolo, L., 2012. First records and community pattern of Arcellinida inhabiting a pristine and remote island from Southeastern Pacific. Chile. Acta Protozool. 51 (2).

Baumgartner, M., Yapi, A., Gröbner-Ferreira, R., Stetter, K.O., 2003. Cultivation and properties of *Echinamoeba thermarum n. sp.*, an extremely thermophilic amoeba thriving in hot springs. Extremophiles 7, 267–274.

Blandenier, Q., Lara, E., Mitchell, E.A., Alcantara, D.M., Siemensma, F.J., Todorov, M., Lahr, D.J., 2017. Nad9/nad7 (mitochondrial nicotinamide adenine dinucleotide dehydrogenase gene)—a new "holy grail" phylogenetic and DNA-barcoding marker for Arcellinida (Amoebozoa)? Eur. J. Protistol. 58, 175–186.

Bobrov, A., Mazei, Y., 2004. Morphological variability of testate amoebae (Rhizopoda: Testacealobosea and Testaceafilosea) in natural populations. Acta Protozool. 43 (2), 133–146.

Burdman, L., Mataloni, G., Mitchell, E.A., Lara, E., 2021. A reassessment of testate amoebae diversity in Tierra del Fuego peatlands: Implications for large scale inferences. Eur. J. Protistol. 80, 125806.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25 (15), 1972–1973.

Duckert, C., Blandenier, Q., Kupferschmid, F.A., Kosakyan, A., Mitchell, E.A., Lara, E., Singer, D., 2018. En garde! redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta gen. nov*. Eur. J. Protistol. 66, 156–165.
Dumack, K., Kahlich, C., Lahr, D.J., Bonkowski, M., 2019. Reinvestigation of *Phryganella paradoxa* (Arcellinida, Amoebozoa) Penard 1902. J. Eukaryot. Microbiol. 66 (2), 232–243.
Ehreberg, C., 1838. Die Infusionsthierchen als vollkommene Organismen. Leopold Voss, Leipzig, Germany.
Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39 (suppl_2), W29–W37.
Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech, 3, 294–299.
Gomaa, F., Todorov, M., Heger, T.J., Mitchell, E.A., Lara, E., 2012. SSU rRNA phylogeny of Arcellinida (Amoebozoa) reveals that the largest Arcellinid genus, *Difflugia* Leclerc 1815, is not monophyletic. Protist 163 (3), 389–399.
Gomaa, F., Kosakyan, A., Heger, T.J., Corsaro, D., Mitchell, E.A., Lara, E., 2014. One alga to rule them all: unrelated mixotrophic testate amoebae (Amoebozoa, Rhizaria and Stramenopiles) share the same symbiont (Trebouxiophyceae). Protist 165 (2), 161–176.
Gomaa, F., Yang, J., Mitchell, E.A., Zhang, W.-J., Yu, Z., Todorov, M., Lara, E., 2015. Morphological and molecular diversification of asian endemic *Difflugia tuberspinifera* (Amoebozoa, Arcellinida): a case of fast morphological evolution in protists? Protist 166 (1), 122–130.
González-Miguéns, R., Soler-Zamora, C., Villar-Depablo, M., Todorov, M., Lara, E., 2022a. Multiple convergences in the evolutionary history of the testate amoeba family Arcellidae (Amoebozoa: Arcellinida: Sphaerothecina): when the ecology rules the morphology. Zool. J. Linn. Soc. 194 (4), 1044–1071.
González-Miguéns, R., Todorov, M., Blandenier, Q., Duckert, C., Porfirio-Sousa, A.L., Ribeiro, G.M., Ramos, D., Lahr, D.J., Buckley, D., Lara, E., 2022b. Deconstructing Difflugia: The tangled evolution of lobose testate amoebae shells (Amoebozoa: Arcellinida) illustrates the importance of convergent evolution in protist phylogeny. Mol. Phylogenet. Evol. 175, 107557.
González-Miguéns, R., Cano, E., Guillen-Oterino, A., Quesada, A., Lahr, D.J., Tenorio-Rodriguez, D., de Salvador-Velasco, D., Velazquez, D., Carrasco-Braganza, M.I., Patterson, T.R., et al., 2023. A needle in a haystack: a new metabarcoding approach to survey diversity at the species level of Arcellinida (Amoebozoa: Tubulinea). Mol. Ecol. Resour. 23 (5), 1034–1049.
Gouy, M., Guindon, S., Gascuel, O., 2010. Seaview version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27 (2), 221–224.
Heger, T.J., Mitchell, E.A., Leander, B.S., 2013. Holarctic phylogeography of the testate amoeba *Hyalosphenia papilio* (Amoebozoa: Arcellinida) reveals extensive genetic diversity explained more by environment than dispersal limitation. Mol. Ecol. 22 (20), 5172–5184.
Jones, F.C., 2008. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environ. Rev. 16 (NA), 45–69.
Katoh, K., Misawa, K., Kuma, K.-I., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30 (14), 3059–3066.
Kosakyan, A., Heger, T.J., Leander, B.S., Todorov, M., Mitchell, E.A., Lara, E., 2012. COI barcoding of Nebelid testate amoebae (Amoebozoa: Arcellinida): extensive cryptic diversity and redefinition of the Hyalospheniidae Schultze. Protist 163 (3), 415–434.
Kosakyan, A., Gomaa, F., Lara, E., Lahr, D.J., 2016. Current and future perspectives on the systematics, taxonomy and nomenclature of testate amoebae. Eur. J. Protistol. 55, 105–117.
Kostka, M., Lares-Jiménez, L.F., Tyml, T., Dyková, I., 2017. *Copromyxa laresi n. sp*. (Amoebozoa:Tubulinea) and transfer of *Cashia limacoides* (page, 1967) to *Copromyx*a Zopf, 1885. J. Eukaryot. Microbiol. 64 (2), 173–182.
Lahr, D.J., Bergmann, P.J., Lopes, S.G., 2008. Taxonomic identity in microbial eukaryotes: a practical approach using the testate amoeba *Centropyxis* to resolve conflicts between old and new taxonomic descriptions. J. Eukaryot. Microbiol. 55 (5), 409–416.
Lahr, D.J., Grant, J., Nguyen, T., Lin, J.H., Katz, L.A., 2011. Comprehensive phylogenetic reconstruction of amoebozoa based on concatenated analyses of SSU-rDNA and actin genes. PLoS One 6 (7), e22780.
Lahr, D.J., Lara, E., Mitchell, E.A., 2012. Time to regulate microbial Eukaryote nomenclature. Biol. J. Linn. Soc. 107 (3), 469–476.
Lahr, D.J., Laughinghouse IV, H.D., Oliverio, A.M., Gao, F., Katz, L.A., 2014. How discordant morphological and molecular evolution among microorganisms can revise our notions of biodiversity on Earth. Bioessays 36 (10), 950–959.
Lahr, D.J., Kosakyan, A., Lara, E., Mitchell, E.A., Morais, L., Porfirio-Sousa, A.L., Ribeiro, G.M., Tice, A.K., Pánek, T., Kang, S., et al., 2019. Phylogenomics and morphological reconstruction of Arcellinida testate amoebae highlight diversity of microbial eukaryotes in the Neoproterozoic. Curr. Biol. 29 (6), 991–1001.
Lara, E., Heger, T.J., Ekelund, F., Lamentowicz, M., Mitchell, E.A., 2008. Ribosomal RNA genes challenge the monophyly of the Hyalospheniidae (Amoebozoa: Arcellinida). Protist 159 (2), 165–176.
Leakey, R. and Lewin, R. 1996. The sixth extinction: Biodiversity and Its survival. A Phoenix paperback.Phoenix.
Macumber, A.L., Blandenier, Q., Todorov, M., Duckert, C., Lara, E., Lahr, D.J., Mitchell, E.A., Roe, H.M., 2020. Phylogenetic divergence within the Arcellinida (Amoebozoa) is congruent with test size and metabolism type. Eur. J. Protistol. 72, 125645.
Marcisz, K., Jassey, V. E., Kosakyan, A., Krashevska, V., Lahr, D. J., Lara, E., Lamentowicz, Ł., Lamen- towicz, M., Macumber, A., Mazei, Y., et al. 2020. Testate amoeba functional traits and their use in Paleoecology. Front. Ecol. Evol., 8, 575966.
Medioli, F., Scott, D.B., Abbott, B., 1987. A case study of protozoan intraclonal variability; taxonomic implications. J. Foraminiferal Res. 17 (1), 28–47.
Mulot, M., Marcisz, K., Grandgirard, L., Lara, E., Kosakyan, A., Robroek, B.J., Lamentowicz, M., Payne, R.J., Mitchell, E.A., 2017. Genetic determinism vs. phenotypic plasticity in protist morphology. J. Eukaryot. Microbiol. 64 (6), 729–739.
Nasser, N.A., Patterson, R.T., Roe, H.M., Galloway, J.M., Falck, H., Sanei, H., 2020. Use of Arcellinida (testate lobose amoebae) arsenic tolerance limits as a novel tool for biomonitoring arsenic contamination in lakes. Ecol. Ind. 113, 106177.
Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32 (1), 268–274.
Pawlowski, J., Holzmann, M., Tyszka, J., 2013. New supraordinal classification of Foraminifera: Molecules meet morphology. Mar. Micropaleontol. 100, 1–10.
Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., Sandberg, R., 2014. Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. 9 (1), 171.
Porfírio-Sousa, A.L., Ribeiro, G.M., Lahr, D.J., 2017. Morphometric and genetic analysis of *Arcella intermedia* and *Arcella intermedia laevis* (Amoebozoa, Arcellinida) illuminate phenotypic plasticity in microbial eukaryotes. Eur. J. Protistol. 58, 187–194.
Ribeiro, G.M., Prado, P.I., Coutinho, R.M., Rillo, M.C., Pereira Junior, S., Porfírio-Sousa, A.L., Lahr, D.J., 2019. Growth rate modulation enables coexistence in a competitive exclusion scenario between microbial eukaryotes. Acta Protozool. 58 (4), 217–233.
Ribeiro, G.M., Porfírio-Sousa, A.L., Maurer-Alcalá, X.X., Katz, L.A., Lahr, D.J., 2020. De novo sequencing, assembly, and annotation of the transcriptome for the free-living testate amoeba *Arcella intermedia*. J. Eukaryot. Microbiol. 67 (3), 383–392.
Roe, H.M., Patterson, R.T., Swindles, G.T., 2010. Controls on the contemporary distribution of lake thecamoebians (testate amoebae) within the greater Toronto area and their potential as water quality indicators. J. Paleolimnol. 43 (4), 955–975.
Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH image to ImageJ: 25 years of image analysis. Nat. Methods 9 (7), 671–675.
Schulz, G., Maraun, M., Völcker, E., Scheu, S., Krashevska, V., 2018. Evaluation of morphological characteristics to delineate taxa of the genus *Trigonopyxis* (Amoebozoa, Arcellinida). Protist 169 (2), 190–205.
Soler-Zamora, C., González-Miguéns, R., Guillén-Oterino, A., Lara, E., 2021. Arcellinida testate amoebae as climate miner's canaries in Southern Spain. Eur. J. Protistol. 81, 125828.
Souto, M.S., Gonçalves, V., Pontevedra-Pombal, X., Raposeiro, P.M., 2021. Distribution of testate amoebae in bryophyte communities in São Miguel Island (Azores Archipelago). Biodivers. Data J. 9.
Todorov, M., Bankov, N., 2019. An atlas of Sphagnum-dwelling testate amoebae in Bulgaria. Advanced Books 1, e38685.
Tsyganov, A. and Mazei, Yu. (2006). Morphology and biometry of *Arcella intermedia* (Deflandre, 1928) comb. nov.. from Russia and a review of hemispheric species of the genus *Arcella* (Testcealobosea, Arcellinida). Protistol. 4(4):361–369.
Wiley, E.O., 1978. The evolutionary species concept reconsidered. Syst. Zool. 27 (1), 17–26.